

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2002, 14:50:01; Search time 11 Seconds

(without alignments)  
867.232 Million cell updates/sec

Title: US-09-856-320A-2\_COPY\_53\_282

Perfect score: 1263  
Sequence: 1 RIINGFECKPHSQPQAAALF.....GVYTKCKYVDWIDETMKNN 230

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1263	100.0	250 1	KLKB_HUMAN
2	739	58.5	250 1	KLK9_HUMAN
3	689	54.6	260 1	NRPN_RAT
4	686	54.3	260 1	NRPN_MOUSE
5	684	54.2	256 1	KLKF_HUMAN
6	683	54.1	260 1	KLK8_HUMAN
7	677.5	53.6	277 1	KLKD_HUMAN
8	649.5	51.4	293 1	KLK5_HUMAN
9	623	49.3	251 1	KLKE_HUMAN
10	610.5	48.3	261 1	KLK1_RAT
11	608	48.1	248 1	KLKQ_HUMAN
12	605.5	47.9	261 1	KLK7_RAT
13	600.5	47.5	261 1	KLK3_MOUSE
14	596.5	47.2	263 1	KLKR_PRANA
15	596	47.2	248 1	TRY3_CHICK
16	593	47.0	246 1	TRY2_RAT
17	590.5	46.8	261 1	KLK8_RAT
18	590	46.7	238 1	TRY3_SALSA
19	589	46.6	246 1	TRY1_RAT
20	586.5	46.4	259 1	KLKC_RAT
21	582.5	46.1	261 1	KLK2_HUMAN
22	582.5	46.1	261 1	KLK9_HUMAN
23	579	45.8	231 1	TRYP_PIG
24	579	45.8	244 1	KLK6_HUMAN
25	575.5	45.6	246 1	KLK_PIG
26	573	45.4	246 1	TRY2_MOUSE
27	572.5	45.3	259 1	KLK2_RAT
28	571.5	45.2	261 1	KLK1_MOUSE
29	570	45.1	247 1	TRY2_CANFA
30	569.5	45.1	244 1	KLKA_RAT
31	569.5	45.1	261 1	KLK6_MOUSE
32	567	44.9	262 1	KLK1_HUMAN
33	566	44.8	243 1	TRY1_BOVIN

34	563	44.6	248 1	TRY1_CHICK	Q90627 gallus gall
35	561.5	44.5	257 1	KLK1_MACFA	Q00726 macaca fasc
36	561.5	44.5	261 1	KLKB_MOUSE	P15946 mus musculu
37	559	44.3	244 1	TRY2_XENLA	P70059 xenopus lae
38	559	44.3	248 1	TRY2_CHICK	Q90628 gallus gall
39	558	44.2	243 1	TRY1_XENLA	P17799 xenopus lae
40	558	44.2	247 1	TRY2_BOVIN	Q29463 bos taurus
41	556.5	44.1	261 1	KLK3_HUMAN	P07288 homo sapien
42	555	43.9	247 1	TRY3_RAT	P08426 rattus norv
43	554.5	43.9	231 1	TRY2_SALSA	P35032 salmo salar
44	554.5	43.9	247 1	TRY4_RAT	P12788 rattus norv
45	554	43.9	258 1	KLK1_PAPHA	Q28773 papio hamad

## ALIGNMENTS

RESULT 1  
KLKB\_HUMAN STANDARD: PRT; 250 AA.  
AC Q9UBX7; 075837; Q9NS65;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Kallikrein 11 precursor (EC 3.4.21.-) (Hippostasin) (Trypsin-like  
DE protease).  
GN KLK11 OR PRSS20 OR TLSP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Hippocampus;  
RX MEDLINE=98438738; PubMed=9765601;  
RA Yoshida S., Taniguchi M., Suemoto T., Oka T., He X.P., Shiosaka S.;  
RT "cDNA cloning and expression of a novel serine protease, TLSP.";  
RL Biochim. Biophys. Acta 1399:225-228(1998).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC TISSUE=Hippocampus, and prostate;  
RX MEDLINE=20329229; PubMed=10872828;  
RA Mitsui S., Yamada T., Okui A., Komitani K., Uemura H., Yamaguchi N.;  
RT "A novel isoform of a kallikrein-like protease, TLSP/hippostasin,  
RT (PRSS20), is expressed in the human brain and prostate.";  
RL Biochem. Biophys. Res. Commun. 272:205-211(2000).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=20130117; PubMed=10662548;  
RA Yousef G.M., Scorilas A., Diamandis E.P.;  
RT "Genomic organization, mapping, tissue expression, and hormonal  
RT regulation of trypsin-like serine protease (TLSP PRSS20), a new  
RT member of the human kallikrein gene family.";  
RL Genomics 63:88-96(2000).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=20510030; PubMed=11054574;  
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McQuay J.,  
RT "Sequencing and expression analysis of the serine protease gene  
RT cluster located in chromosome 19q13 region.";  
RL Gene 257:119-130(2000).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RA Iamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,  
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Gaines J.,  
RA Dangnan L., Erlor A., Christensen M., Georgescu A., Avila J., Liu S.,  
RA Andrease T., Frankheim M., Attix C., Amico-Keller G., Coefficient J.,  
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,  
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
RA Olsen A.S., Carrano A.V.;  
RT "Sequence analysis of chromosome 19q13.4.";

```

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE-Testis;
RA Strauberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- PUNCTION: POSSIBLE MULTIFUNCTIONAL PROTEASE. EFFICIENTLY CLEAVES
CC B2-PHE-ARG-4-METHYLCOMANTYL-7-AMIDE, A KALLIKREIN SUBSTRATE, AND
CC WEAKLY CLEAVES OTHER SUBSTRATES FOR KALLIKREIN AND TRYPSIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, SKIN AND PROSTATE. ISOFORM
CC 1 IS EXPRESSED PREFERENTIALLY IN BRAIN; ISOFORM 2 IN PROSTATE.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB012917; BAA33404.1; ALT_INT.
DR EMBL: AB013730; BAA8713.1; -.
DR EMBL: AB041036; BAA96797.1; -.
DR EMBL: AF164623; AAD47815.1; -.
DR EMBL: AF243527; AAG33364.1; -.
DR EMBL: AC011473; AAG23257.1; -.
DR EMBL: BC022068; AAR22068.1; -.
DR HSSP: P00763; IDPO.
DR MEROPS: S01.257; -.
DR Genew: HGNC:6359; KLK11.
DR MIM: 604434; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYPSIN_DOM. 1.
DR PROSITE: PS50240; TRYPSIN_DOM. 1.
DR PROSITE: PS00134; TRYPSIN_HIS. 1.
DR PROSITE: PS00135; TRYPSIN_SER. 1.
KW Hydrolyase; Serine protease; Glycoprotein; Signal; Zymogen;
KW Alternative splicing.
FT SIGNAL 1
FT PROPEP 19 21 POTENTIAL.
FT CHAIN 22 250 ACTIVATION PEPTIDE (POTENTIAL).
FT ACT_SITE 62 62 KALLIKREIN 11.
FT ACT_SITE 110 110 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 203 203 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULEID 28 163 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULEID 47 63 BY SIMILARITY.
FT DISULEID 135 237 BY SIMILARITY.
FT DISULEID 142 209 BY SIMILARITY.
FT DISULEID 174 188 BY SIMILARITY.
FT DISULEID 199 224 BY SIMILARITY.
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1 1 M -> MORLUMLDMWSSGGLTAAKEPCARSSPLDM
FT (IN ISOFORM 2).
SQ SEQUENCE 250 AA; 27466 MW; 192D910BBDCD7A56 CRC64;
Query Match 100.0%; Score 1263; DB 1; Length 250;
Best Local Similarity 100.0%; Pred. No. 7, 1e-109;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 61 ECCEQTATATSPHPGPNNSLPKKDRNDIMLVKMASPVSTWAVRPLTSSRCVACT 120
DB 81 ECCEQTATATSPHPGPNNSLPKKDRNDIMLVKMASPVSTWAVRPLTSSRCVACT 140
QY 121 SCLISGMSSTSPQLRPHTRCANITITIEHOKENAYPGNITPTWCASVQEGKSCQ 180
DB 141 SCLISGMSSTSPQLRPHTRCANITITIEHOKENAYPGNITPTWCASVQEGKSCQ 200
QY 181 GDSGGPLVCNQSLOGIISWGDPCATIRKPGVYTKYCKYVDWIQETKNN 230
DB 201 GDSGGPLVCNQSLOGIISWGDPCATIRKPGVYTKYCKYVDWIQETKNN 250

RESULT 2
ID KLK9_HUMAN STANDARD; PRT; 250 AA.
AC 09UK09;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Kallikrein 9 precursor (EC 3.4.21.-) (Kallikrein-1-like protein 3) (KLK-
DE L3).
GN KLK9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20118156; PubMed=10652563;
RA Yousef G.M., Luo L.-Y., Diamandis E.P.;
RT "Identification of novel human kallikrein-1-like genes on chromosome
RT 19q13.3-q13.4."
RL Anticancer Res. 19:2843-2852(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20247258; PubMed=10783266;
RA Yousef G.M., Diamandis E.P.;
RT "The expanded human kallikrein gene family: locus characterization and
RT molecular cloning of a new member, KLK-13."
RL Genomics 65:184-194(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20510030; PubMed=11054574;
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
RA Moss P., Paepel B., Wang K.;
RT "Sequencing and expression analysis of the serine protease gene
RT cluster located in chromosome 19q13 region."
RL Gene 257:119-130(2000).
RN [4]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Wiswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Ganes J.,
RA Danganan L., Ertler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Duarte S., Lucas S., Bruce R., Attix C., Amico-Keller G., Coefield J.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carraro A.V.;
RT "Sequence analysis of chromosome 19q13.4."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- TISSUE SPECIFICITY: SKIN, THYMUS, TRACHEA, CEREBELLUM AND SPINAL
CC CORD.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

CC EMBL; AF135026; AAD36427.2; -  
DR EMBL; AF243527; AAC33362.1; -  
DR EMBL; AC011473; AAG23255.1; -  
DR HSSP; P00763; IDPO.  
DR MEROPS; S01.307; -  
DR GeneW; HGNC:6370; KIK9.  
DR MIM; 605504; -  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00089; trypsin\_1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_spec; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KM Hydrolyase; Serine protease; Glycoprotein; Signal.  
FT SIGNAL 1  
FT CHAIN 16 250  
FT ACT\_SITE 63 63 KALLIKREIN 9.  
FT ACT\_SITE 111 111 CHANGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 204 204 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 29 164 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 48 64 BY SIMILARITY.  
FT DISULFID 136 238 BY SIMILARITY.  
FT DISULFID 143 210 BY SIMILARITY.  
FT DISULFID 175 189 BY SIMILARITY.  
FT CARBOHYD 200 225 BY SIMILARITY.  
FT CARBOHYD 131 131 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 166 166 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 211 211 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 230 AA; 27512 MW; F2785245B03E98B CRC64.

Query Match 58.5%; Score 739; DB 1; Length 250;  
Best Local Similarity 58.1%; Pred. No. 8,8e-61;  
Matches 133; Conservative 35; Mismatches 61; Indels 0; Gaps 0

QY 1 RIRKGECKRPHSQMOALFEKTRLLCGATLTAPRWMLTAHCLKRPRTVHLGHNLQKE 60  
+ + + + + | : : : : : | + + + + + | : : : : : | + + + + + |  
Db 22 RAIGAEBCRPNSQPDAQLFLITRFECGATLTSIDRMLLAACHCRPYLMVRGEGHLMWM 81

QY 61 EGCEGTFRATESFPRLPGFNNSLPNDKHNDIMLVMAASPVSTVMVRPLTISRCTAGT 120  
+ + + + + | : : : : : | + + + + + | : : : : : | : : : : : |  
Db 82 EGPEDLFRYTDFPPHPGNKKDLANDHDDIMLRPROARLSPAVQPLNLSQTCSPEM 141

QY 121 SCLISGMGSTSPOLRLPHTLRCANITTEIEHKCENAPNTIDPMVCASVOEGGDSGO 180  
+ + + + + | : : : : : | + + + + + | : : : : : | : : : : : |  
Db 142 QCLISGMGCVASSPKALFPTVLQCANISTLENELCHMAVPGHISDMLCAGMEGGSGSO 201

QY 181 GDSEGPLVCNOSLOGIISMGDPCALTNRKPGYTYRVCKYDVMIQETMKN 229  
+ + + + + | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 202 GDSGEPVLCNGTLAGVSGAIPCSRPRPRRAYTVSVCHYLWIQEIEMEN 250

RESULT 3  
NRPN\_RAT NRPN\_RAT STANDARD: PRT: 260 AA.

AC 088780;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Neuropeptide precursor (EC 3.4.21.-) (NP) (Kallikrein 8) (Brain serine  
protease 1)  
GN KIK8 OR PRSS19 OR NRPN OR BSP1.  
OC Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_Taxid=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Fischer; TISSUE=Brain;  
FX MEDLINE=96389725; PubMed=9722524;  
NA Davies B.J., Pickard B.S., Steel M., Morris R.G.M., Lathie R.;

[illegible]



RT "A novel serine proteinase-like sequence from human brain.";  
 RL Biochim. Biophys. Acta 1218:225-228(1994).  
 CC -1- FUNCTION: Protease whose physiological substrate is not yet known.  
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2, 3 and 4; are  
 CC produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: Highest expression in the thyroid gland. Also  
 CC expressed in the prostate, salivary, and adrenal glands and in the  
 CC colon testis and kidney.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF242195; AAG09469.1; -;  
 DR EMBL: AF242195; AAG09470.1; -;  
 DR EMBL: AF242195; AAG09471.1; -;  
 DR EMBL: AF242195; AAG09472.1; -;  
 DR EMBL: AF243527; AAG33354.1; -;  
 DR EMBL: X75363; CAAS3145.1; ALT\_SEQ.  
 DR HSSP: P00763; IDPO.  
 DR MEROPS: S01.081; -;  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR001254; Ser.protease\_Try.  
 DR Pfam: PF00089; trypsin.1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR SMART: SM00020; TRY\_PSPC.1.  
 DR PROSITE: PS00240; TRYPSIN\_DOM.1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS.1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; FALSE\_NEG.  
 KM Hydroxylase; Serine protease; Glycoprotein; Signal; Zymogen;  
 KW Alternative splicing.  
 FT SIGNAL 1 16  
 FT PROPEP 17 21  
 FT CHAIN 22 256  
 FT ACT\_SITE 62 62  
 FT ACT\_SITE 106 106  
 FT ACT\_SITE 209 209  
 FT CARBOHYD 171 171  
 FT CARBOHYD 232 232  
 FT CARBOHYD 122 206  
 FT VARSPPLIC 161 161  
 FT VARSPPLIC 162 256  
 FT VARSPPLIC 147 160  
 FT CONFLICT 147 160  
 SQ SEQUENCE 256 AA; 28087 MW; B5EBF8D6022786B5 CRC64;  
 Query Match 54.2%; Score 684; DB 1; Length 256;  
 Best Local Similarity 51.7%; Pred. No. 1e-55; Indels 14; Gaps 3;  
 Matches 124; Conservative 37; Mismatches 65;  
 QY 1 RIKEGECKHSPQWQAALEKTRKLGATLAPRLMLTAHCLKPRYIVHLGQHNLOKE 60  
 Db 21 KLEEGECACAPHSQPMQVALTERGRFNGCAGSLISPHWLTSAHQSRMRRLDEHNKRK 80  
 QY 61 EGCEQRTATSEPPHGFQFNNSLKNKDRNDIMLKASPVSTWAVRPLTLSSRCVYAGT 120  
 Db 81 DGEFOETSRVLPHPRYE---ARSHRNDIMLRVQPARLNPOVPAVLPTRCPHPGE 136  
 QY 121 SCISIGMGSTSP---QLRLPHRLRCANTIIIEHCKEENAYPGNIDTWVCAS 170  
 Db 137 ACVAVSGGLVSHNEPGTAGSPRSQVSLPDLHCANISLISDTSQDSYSPERLNTWVCAG 196  
 QY 171 VDEGKSDSCGDSGGPLVQNSLOGIISWGDCPAITRKGVYTKYCKYDQIETKKN 230  
 Db 197 AEGRAESCEGDSGGLVCGGILQGLIVSMGDVPCDNTKRGVYTKYCHYLEWRETKRN 256

RESULT 6  
 KLR8\_HUMAN STANDARD; PRT; 260 AA.  
 AC 060259; Q9U047; Q9HC83; Q9UL19;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Neuropilin precursor (EC 3.4.21.-) (NP) (Kallikrein 8) (Ovasin) (Serine  
 DE protease TADG-14) (Tumor-associated differentially expressed gene-14  
 DE protein).  
 GN KLR8 OR PRSS19 OR TADG14 OR NRPN.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Hippocampus;  
 RX MEDLINE=98372070; PubMed=9714609;  
 RA Yoshida S., Taniguchi M., Hirata A., Shiosaka S.;  
 RT "Sequence analysis and expression of human neuropilin cDNA and gene.";  
 RL gene 213:9-16(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE=Brain;  
 RX MEDLINE=99203457; PubMed=10102990;  
 RA Mitsui S., Tsuruoka N., Yamashiro K., Nakazato H., Yamaguchi N.;  
 RT "A novel form of human neuropilin, a brain-related serine protease, is  
 RT generated by alternative splicing and is expressed preferentially in  
 RT human adult brain.";  
 RL Eur. J. Biochem. 260:627-634(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Ovary;  
 RX MEDLINE=99413504; PubMed=10485494;  
 RA Underwood L.J., Tanimoto H., Wang Y., Shigemasa K., Parmley T.H.,  
 RA O'Brien T.J.;  
 RT "Cloning of tumor-associated differentially expressed gene-14, a novel  
 RT serine protease overexpressed by ovarian carcinoma.";  
 RL Cancer Res. 59:4435-4439(1999).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Gan L., Gellinas R., Gown A.M., Moss P., Smith R., Wang K.;  
 RT "Molecular cloning and characterization of a novel serine protease,  
 RT ovasin, a potential molecular marker for ovarian carcinomas.";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=20510030; PubMed=11054574;  
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
 RA Moss P., Paepert B., Wang K.;  
 RT "Sequencing and expression analysis of the serine protease gene  
 RT cluster located in chromosome 19q13 region.";  
 RL gene 257:119-130(2000).  
 RN [6]  
 RP SEQUENCE OF 1-164 FROM N.A. (ISOFORM 1).  
 RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stiivaen S.,  
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Gaines J.,  
 RA Danganan L., Frankheim M., Attix C., Amico-Keller G., Coefield J.,  
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,  
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
 RA Olsen A.S., Carrano A.V.;  
 RT "Sequence analysis of chromosome 19q13.4.";  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: SUGGESTED TO BE INVOLVED IN KINDLING EPILEPTOGENESIS AND  
 CC HIPPOCAMPAL PLASTICITY.  
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: ISOFORM 1 IS PREDOMINANTLY EXPRESSED IN THE

CC PANCREAS WHITE ISOFORM 2 IS EXPRESSED IN ADULT BRAIN AND  
 CC HIPPOCAMPUS. BOTH FORMS ARE ALSO FOUND IN FETAL BRAIN AND  
 CC PLACENTA. NOT DETECTED IN KIDNEY, SPLEEN, LIVER AND LUNG.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, KALLIKREIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; AB009849; BAA28673.1; -  
 CC EMBL; AB012761; BAA28676.1; -  
 CC EMBL; AB010780; BAA86884.1; -  
 CC EMBL; AB008390; BAA82665.1; -  
 CC EMBL; AB008927; BAA82666.1; -  
 CC EMBL; AF055982; AAD56050.1; -  
 CC EMBL; AF095742; AAD5979.1; -  
 CC EMBL; AF095743; AAD29574.1; -  
 CC EMBL; AF243527; AAG33361.1; -  
 CC EMBL; AC011473; AAG32254.1; -  
 CC HSSP; 061955; INPM.  
 CC MEROPS; S01.244; -  
 CC Gene; HGNC:6369; KLK8.  
 CC MIM; 605644; -  
 CC InterPro; IPR001314; Chymotrypsin.  
 CC InterPro; IPR001254; Ser\_protease\_Try.  
 CC Pfam; PF00089; trypsin.1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC SMART; SM00020; TRY-SPEC.1.  
 CC PROSITE; PS02040; TRYPSIN\_DOM.1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS.1.  
 CC PROSITE; PS00135; TRYPSIN\_SER.1.  
 CC HydroLase; Serine protease; Glycoprotein; zymogen; signal;  
 CC Alternative splicing.  
 CC FT SIGNAL 1 28  
 CC FT PROPE 29 32 BY SIMILARITY.  
 CC FT CHAIN 33 260 NEUROPSIN.  
 CC FT ACT\_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 CC FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 CC FT ACT\_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 CC FT DISULFID 39 173 BY SIMILARITY.  
 CC FT DISULFID 58 74 BY SIMILARITY.  
 CC FT DISULFID 145 246 BY SIMILARITY.  
 CC FT DISULFID 152 218 BY SIMILARITY.  
 CC FT DISULFID 184 198 BY SIMILARITY.  
 CC FT DISULFID 208 233 BY SIMILARITY.  
 CC FT CARBOHYD 110 110 N-LINKED (GLCNAC...) (POTENTIAL).  
 CC FT VARSPLIC 23 23 A -> AACGSLDLTKLYAENLPCVHLNPMQWSPQSHCPRG  
 CC FT WRSNPLPPAA (IN ISOFORM 2).  
 CC SEQUENCE 260 AA; 28048 MW; EF439E5B8C3B660 CRC64;  
 SQ  
 Query Match 54.1%; Score 683; DB 1: Length 260;  
 Best Local Similarity 51.6%; Pred. No. 1.3e-55;  
 Matches 116; Conservative 43; Mismatches 64; Indels 2; Gaps 2;

RESULT 7  
 ID KIKD\_HUMAN STANDARD; PRT; 277 AA.  
 AC Q9UKR3; Q9Y433;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Kallikrein 13 precursor (EC 3.4.21.-) (Kallikrein-like protein 4)  
 DE (KIK-14).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP MEDLINE-20229789; PubMed-10766816;  
 RX Youssef G.M., Chang A., Diamandis E.P.;  
 RA "Identification and characterization of KLK-14, a new kallikrein-like  
 RT gene that appears to be down-regulated in breast cancer tissues.";  
 RL J. Biol. Chem. 275:11891-11898(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,  
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,  
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Gaines J.,  
 RA Dangnan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,  
 RA Duarte T., Frankel M., Altix C., Amico-Keller G., Coiffield J.,  
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,  
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
 RA Olsen A.S., Carrano A.V.;  
 RT "Sequence analysis of chromosome 19q13.4.";  
 RL Submitted (Oct-2000) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-180 FROM N.A.  
 RC TISSUE=uterus;  
 RA Ansoorge W., Wilkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;  
 RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN PROSTATE, BLAST, TESTIS AND  
 CC SALIVARY GLAND.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, KALLIKREIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; AF135024; AAD26425.2; -  
 CC EMBL; AC011473; AAG33259.1; -  
 CC EMBL; AL050220; CAB3320.1; ALT\_INIT.  
 CC HSSP; P00763; IDPO.  
 CC MEROPS; S01.306; -  
 CC Gene; HGNC:6361; KLK13.  
 CC MIM; 605505; -  
 CC InterPro; IPR001314; Chymotrypsin.  
 CC InterPro; IPR001254; Ser\_protease\_Try.  
 CC Pfam; PF00089; trypsin.1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC SMART; SM00020; TRY-SPEC.1.  
 CC PROSITE; PS02040; TRYPSIN\_DOM.1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS.1.  
 CC PROSITE; PS00135; TRYPSIN\_SER.1.  
 CC HydroLase; Serine protease; Glycoprotein; signal;  
 CC FT SIGNAL 1 16  
 CC FT CHAIN 17 277 KALLIKREIN 13.  
 CC FT ACT\_SITE 76 76 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 CC FT ACT\_SITE 124 124 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT

```

FT ACT_SITE 218 218 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 42 178 BY SIMILARITY.
FT DISULFID 61 77 BY SIMILARITY.
FT DISULFID 157 224 BY SIMILARITY.
FT DISULFID 189 203 BY SIMILARITY.
FT DISULFID 214 239 BY SIMILARITY.
FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 170 180 UNPKLOCAN -> GMPHPRMEAP (IN REF. 3).
SQ SEQUENCE 277 AA: 30570 MW: BA8A9EBDCFB5D542 CRC64;

Query Match 53.6%; Score 677.5; DB 1; Length 277;
Best Local Similarity 53.3%; Pred. No. 4.3e-55;
Matches 120; Conservative 41; Mismatches 63; Indels 1; Gaps 1;

OY 5 GFCEKPHSQPQWQAALFEKTRILCGATLLAPRLTLTAHCLKPRYIVHILGOHLOKEGCE 64
DB 39 GYCFPHSQPQWQAALLVQGLLGGVLYVHKRWVLTAAHCLKEGLKYLGLHAGLGRVAGE 98
OY 65 QTRATESFPHPGFNNSLPKNDHNDIMLVKMASPVSTWAVRPLTSL-SRCYTAGTSCSL 123
DB 99 QVREVVHSHIPREYRSPRTILNHDHIMLELQSPVQLFGYIQLPLSHNNRLTPGTTCR 158
OY 124 ISGMSGSSPOLRPHTLRCANTITIEHOKCENAPGNTIDTWCASVQEGKDSQCGDS 183
DB 159 VSGMGTTPSPQVNYPRKLOCANIQLRSEDECRQYPERKIDNMICAGTKEGSGDSCGDS 218
OY 184 GGPVACNOSLOGIISWGDPCATRRKPGVYTKVKYVDWIOETMKN 228
DB 219 GGPVACNKRITKXGIVSMGDFPCGQPDREGVYTRVSRVYLMIRETIR 263

RESULT 8
KLKE_HUMAN
ID KLKE_HUMAN STANDARD; PRT: 293 AA.
AC 09Y337: 09H8G8.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Kallikrein 5 precursor (BC 3.4.21.-) (Stratum corneum tryptic enzyme)
DE (Kallikrein-like protein 2) (KLK-L2).
GN KLK5 OR SCFE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stratum corneum;
RC MEDLINE=9445563; PubMed=10514489;
RA Bratsand M., Egelrud T.;
RT "Purification, molecular cloning, and expression of a human stratum
RT corneum trypsin-like serine protease with possible function in
RT desquamation."
RL J. Biol. Chem. 274:30033-30040(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=20118156; PubMed=10652563;
RA Yousef G.M., Luo L.-Y., Diamandis E.P.;
RT "Identification of novel human kallikrein-like genes on chromosome
RT 19q13.3-q13.4."
RL Anticancer Res. 19:2843-2852(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=20510030; PubMed=11054574;
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCaig J.,
RA Moss P., Paepel B., Wang K.;
RT "Sequencing and expression analysis of the serine protease gene
RT cluster located in chromosome 19q13 region."
RL Gene 257:119-130(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=ovary;

```

```

RA Strausberg R.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY BE INVOLVED IN DESQUAMATION.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SKIN, BREAST, BRAIN AND TESTIS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF168768; AAF03101.1; -
DR EMBL: AF135028; AAD26429.1; -
DR EMBL: AF243527; AAG33358.1; -
DR EMBL: BC008036; AAB08036.1; -
DR HSSP: P00763; IDPO.
DR MEROPS: S01.017; -.
DR GeneW: HGNC:6366; KLK5.
DR MIM: 605643; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Tryp.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYPSIN_SPC. 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Glycoprotein; Signal.
FT SIGNAL 1 22
FT CHAIN 23 293 KALLIKREIN 5.
FT ACT_SITE 108 108 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 153 153 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 245 245 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 73 206 BY SIMILARITY.
FT DISULFID 93 109 BY SIMILARITY.
FT DISULFID 178 279 BY SIMILARITY.
FT DISULFID 185 251 BY SIMILARITY.
FT DISULFID 217 231 BY SIMILARITY.
FT DISULFID 241 266 BY SIMILARITY.
FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 25 36 MISSING (IN REF. 3).
SQ SEQUENCE 293 AA: 32020 MW: D92C92F5609E5946 CRC64;

Query Match 51.4%; Score 649.5; DB 1; Length 293;
Best Local Similarity 50.9%; Pred. No. 1.7e-52;
Matches 118; Conservative 40; Mismatches 67; Indels 7; Gaps 4;

OY 1 RIIFGECKPHSQPQWQAALFEKTRILCGATLLAPRLTLTAHCLKPRYIVHILGOHLOK 59
DB 66 RIIFGSCDMHTQPMQWQAALLRPNQLYCGAVLVHPQWLLTAHCLKPRYIVHILGOHLSLP 125
OY 60 -EEGCEQTRATESFPHPGFNNSLPKNDHNDIMLVKMASPVSTWAVRPLTSLSSRCYTA 118
DB 126 VYESGQMGVGVSRIPRYS---HPGSHNDMLKLNRRIRPTDVRINVSHPSPA 181
OY 119 GFSCLISGMSGSSPOLRPHTLRCANTITIEHOKCENAPGNTIDTWCASVQEGKDS 178
DB 182 GTGCLVSGMGTTPSPQVNYPRKLOCANIQLRSEDECRQYPERKIDNMICAGTKEGSGDSC 240
OY 179 CGGDSGPPVACNOSLOGIISWGDPCATRRKPGVYTKVKYVDWIOETMKN 230
DB 241 CGGDSGPPVACNOSLOGIISWGDPCARPRPVRVYTMCKFTKWIQETIQAN 292

RESULT 9
KLKE_HUMAN

```







```

RT      expressed in the kidney." ;
RL      J. Biochem. 105:834-840(1989).
[4]
RN      SEQUENCE FROM N.A.
RP      MEDLINE=89214217; PubMed=2708383;
RA      Wines D.R., Brady J.M., Pritchett D.B., Roberts J.L., Macdonald R.J.;
RT      "Organization and expression of the rat kallikrein gene family.";
RL      J. Biol. Chem. 264:7653-7662(1989).
[5]
RN      SEQUENCE OF 48-261 FROM N.A.
RP      MEDLINE=86131678; PubMed=3004582;
RA      Gerald W.L., Chao J., Chao L.;
RT      "Immunological identification of rat tissue kallikrein cDNA and
RT      characterization of the kallikrein gene family.";
RL      Biochim. Biophys. Acta 866:1-14(1986).
CC      -1- FUNCTION: GLANDULAR KALLIKREIN CLEAVE MET-LYS AND ARG-SER BONDS
CC      IN KININOGEN TO RELEASE LYS-BRADYKININ.
CC      -1- CATALYTIC ACTIVITY: Preferential cleavage of Arg-1-Xaa bonds in
CC      small molecule substrates. Highly selective action to release
CC      kallidin (Lysyl-Diarginine) from kininogen involves hydrolysis of
CC      Met-1-Xaa or Leu-1-Xaa.
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, KALLIKREIN SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation --
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb-rib.ch).
CC      -----
DR      EMBL; J00758; -; NOT_ANNOTATED_CDS.
DR      EMBL; M1563; AAA41464.1; ALT_INT.
DR      EMBL; M23876; AAA41462.1; -.
DR      EMBL; M23874; AAA41462.1; JOINED.
DR      EMBL; M23875; AAA41462.1; JOINED.
DR      EMBL; D00448; BAA00346.1; ALT_INT.
DR      EMBL; D00446; BAA00346.1; JOINED.
DR      EMBL; D00447; BAA00346.1; JOINED.
DR      EMBL; X03560; CAA27247.1; -.
DR      PIR; A00944; KORP.
DR      PIR; A23863; A23863.
DR      PIR; JX0073; JX0073.
DR      PIR; A33359; A33359.
DR      HSSP; P00757; 1SGF.
DR      MEROPS; S01.405; -.
DR      InterPro: IPR001314; Chymotrypsin.
DR      InterPro: IPR001254; Ser_protease_Try.
DR      Pfam; PF00089; Trypsin_1.
DR      PRINTS; PR00722; CHYMOTRYPSIN.
DR      SMART; SM00020; TRYD_SPC; 1.
DR      PROSITE; PS00240; TRYPSIN_DOM; 1.
DR      PROSITE; PS00134; TRYPSIN_HIS; 1.
DR      PROSITE; PS00135; TRYPSIN_SER; 1.
KW      Hydroxylase; Serine protease; Glycoprotein; Multigene family; Zymogen;
KW      Signal.
FT      SIGNAL          1      18      PROBABLE.
FT      PROPEP         19      24      ACTIVATION PEPTIDE (PROBABLE).
FT      CHAIN          25      261     GLANDULAR KALLIKREIN 1.
FT      CHAIN          25      111     CHAIN 1.
FT      CHAIN          112      261     CHAIN 2.
FT      ACT_SITE       65      65      CHANGE RELAY SYSTEM.
FT      ACT_SITE       120      120     CHANGE RELAY SYSTEM.
FT      ACT_SITE       213      213     CHANGE RELAY SYSTEM.
FT      DISULFID        31      173     BY SIMILARITY.
FT      DISULFID        50      66      BY SIMILARITY.
FT      DISULFID       152      219     BY SIMILARITY.
FT      DISULFID       184      198     BY SIMILARITY.
FT      DISULFID       209      234     BY SIMILARITY.
FT      CARBOHYD       108      108     N-LINKED (GLCNAC... ) (PROBABLE).
FT      SEQUENCE       261 AA; 28852 MW; 2F299C00227A7882B CMC64;
Query Match          48.3%; Score 610.5; DB 1; Length 261;

```

```

Best Local Similarity: 45.6%, Pred. No. 5.7e-49; Indels    7; Gaps    1;
Matches 108; conservative   43; Mismatches 79;
OY      1 RIINKECKPHSQWQALFETRLICGATLLIAPRWLTAAHCLKPRYLVIHLGHNLQKE 60
       1:: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: :
DB      24 RLVGVNCEMNSNQPMOVAVYYEFGELTGCGLVLDPSWMTITAAICAINDQVMYGRNNIED 83
OY      61 ECGEDTRATESPFPHPGFENSL-----PNKDRNDIMLYKMSPSVSTIAVAERPLTSS 113
       | :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: :
DB      84 EPFAOHRIIVYSOFPPHFPGMODLIMNHTRPGSDYENDMLTLHSOPADITDGVKKVIDLP I 143
OY     114 RCVTAGTSCLIGKMSSTSPOLRHLHTLCARNTIIIEHQEKCNARPNGITDMNCASVQE 173
DB     144 EEPKVGSGTCLAGWSGITPDGLESTDIDQCIVIDLSLNCKVEAKHEEVTDLMTAGEMXD 203
OY     174 GSKDSOCGSQGPVLVCNOSLOGIISMGCPSCAITRKRPGYRVVKCYVDMDPTMKNN 230
DB     204 GKRDICKDGSGGPLCNGVLQGITSMGNFPCGEPAKPPIYTTLIKFTFIKEVKMEN 260

RESULT_11
KLKC_HUMAN
ID      KLKC_HUMAN          STANDARD;              PRO;           248 AA.
AC      O9UKR0; O9UKR1;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Kallikrein 12 precursor (EC 3.4.21.-) (Kallikrein-like protein 5)
        (Klk-15).
GN      KLK12 OR KLK15.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
RX      NCBI_Taxid=9606;
RN      [1]
RS      SEQUENCE FROM N.A. (ISOFORM 1),
RA      MEDLINE=20118156; PubMed=10652563;
RT      Yousef G.M., Luo L.-Y., Diamandis E.P.;
RL      "Identification of novel human kallikrein-like genes on chromosome
         19q13.3-q13.4.";
RN      RL Anticancer Res. 19:2843-2852(1999).
[2]
RN      RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RA      Yousef G.M., Magklara A., Scorilas A., Diamandis E.P.;
RT      "Cloning of new alternatively spliced forms of the kallikrein-like
RL      gene 5 (Klk-15)." ;
RN      Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
[3]
RN      RN SEQUENCE FROM N.A. (ISOFORM 1).
RA      MEDLINE=20510030; PubMed=11054574;
RB      Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McQuaid J.,
RT      Moss P., Paeger B., Wang K.;
RL      "Sequencing and expression analysis of the serine protease gene
         cluster located in chromosome 19q13 region." ;
RN      Gene 257:119-130(2000).
[4]
RN      RN SEQUENCE FROM N.A. (ISOFORM 2).
RA      Lamerdin J.E., MCCready P.M., Skowronski E., Viswanathan V.,
RB      Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stiliwgen S.,
RC      Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Garnes J.,
RD      Dangnan L., Euler A., Christensen M., Georgescu A., Avila J., Liu S.,
RE      Andreise T., Frankheim M., Attix C., Amico-Keller G., Coeffield J.,
RF      Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
RG      Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RH      Olsen A.S., Carraro A.V.;
RI      "Sequence analysis of chromosome 19q13.4." ;
RJ      Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.
RK      -1- SUBCELLULAR LOCATION: Secreted (Probable).
RL      -1- ALTERNATIVE PRODUCTS: 2 isoforms: 1 (shown here) and 2; are
RM      produced by alternative splicing.
RN      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKEIN SUBFAMILY.
CC      This SWISS-PROT entry is copyright It is produced through a collaboration
CD      CC
```



QY 1 RIIFGECKRHSQPMOALFEKTRLLCGATLIPRLMLTAHCLKPRYIVHGOHNLQKE 60  
 DB 24 RVIIGFCEKRNQPMOALFEKTRLLCGATLIPRLMLTAHCLKPRYIVHGOHNLQKE 83  
 QY 61 EGGEOTRTATESPPHGFNNLSL-----PNKDRNDIMLVKASPVSTMAVRPLTSS 113  
 DB 84 EPRNQHVLVSQSPHPHYKPFELMKRNRKRDGDSHNDMLHLISQPDITDGVKVDLPT 143  
 QY 114 RCYTAGTSCILSGWGSSPOLRLPRLTLCANTTIEHOKCENAYPGNTDPMVCAVQOE 173  
 DB 144 EEPKVGSTCLASGWSGTRPKPFTDDICVNLKLNELNECIRAYKEKRYVDIMACAGELE 203  
 QY 174 GGDSCGDSGGPVLVNCQSLGIIISWGDPACATRRPGVYTKYCKYVDIQTETKNN 230  
 DB 204 GGDTCGDSGGPVLVNCQSLGIIISWGDPACATRRPGVYTKYCKYVDIQTETKNN 260

RESULT 13  
 KLR3\_MOUSE STANDARD: PRT: 261 AA.  
 AC P00756;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glandular kallikrein K3 precursor (EC 3.4.21.35) (tissue kallikrein)  
 DE (mGK-3) (7S nerve growth factor gamma chain) (gamma-NGF).  
 GN KLR3 OR KLR-3 OR NGFG.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85076169; PubMed=6548955;  
 RA Ullrich A., Gray A., Wood W.I., Hayflick J., Seeburg P.H.;  
 RT "Isolation of a cDNA clone coding for the gamma-subunit of mouse  
 RT nerve growth factor using a high-stringency selection procedure";  
 RL DNA 3:387-392(1984).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85257431; PubMed=3848399;  
 RA Evans B.A., Richards R.I.;  
 RT "Genes for the alpha and gamma subunits of mouse nerve growth factor  
 RT are contiguous";  
 RL EMBO J. 4:133-138(1985).  
 RN (3)  
 RP SEQUENCE OF 25-261.  
 RX MEDLINE=81264363; PubMed=7263706;  
 RA Thomas K.A., Baglan N.C., Bradshaw R.A.;  
 RT "The amino acid sequence of the gamma-subunit of mouse submaxillary  
 RT gland 7 S nerve growth factor";  
 RL J. Biol. Chem. 256:9156-9166(1981).  
 RN (4)  
 RP X-RAY CRYSTALLOGRAPHY (3.15 ANGSTROMS) OF 7S COMPLEX.  
 RC STRAIN=swiss Webster; TISSUE=Submaxillary gland;  
 RX MEDLINE=98035451; PubMed=9351801;  
 RA Bax B., Blundell T.L., Murray-Rust J., McDonald N.O.;  
 RT "Structure of mouse 7S NGF: a complex of nerve growth factor with  
 RT four binding proteins";  
 RL Structure 5:1275-1285(1997).  
 CC -1- FUNCTION: 7S NGF ALPHA CHAIN STABILIZES THE 7S COMPLEX. THE BETA  
 CC DIMER PROMOTES NEURITE GROWTH. THE GAMMA CHAIN IS AN ARGININE-  
 CC SPECIFIC PROTEASE; IT MAY ALSO HAVE PLASMINOGEN ACTIVATOR  
 CC ACTIVITY, AS WELL AS MITOGENIC ACTIVITY FOR CHICK EMBRYO  
 CC FIBROBLASTS.  
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage of Arg-1-Xaa bonds in  
 CC small molecule substrates. Highly selective action to release  
 CC kallidin (Lysyl-bradykinin) from kininogen involves hydrolysis of  
 CC Met-1-Xaa or Leu-1-Xaa.  
 CC -1- SUBUNIT: 7S NERVE GROWTH FACTOR IS COMPOSED OF TWO ALPHA CHAINS,  
 CC A BETA DIMER COMPOSED OF IDENTICAL CHAINS, AND TWO GAMMA CHAINS.  
 CC -1- MISCELLANEOUS: THIS PRECURSOR IS CLEAVED INTO SEGMENTS TO PRODUCE  
 CC THE ACTIVE FORM OF THE GAMMA CHAIN, WHICH OCCURS NATURALLY AS

CC COMBINATIONS OF EITHER TWO OR THREE SEGMENTS HELD TOGETHER BY  
 CC DISULFIDE BONDS: B1 + A OR B1 + C + B2.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, KALLIKREIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: X01369; CAA25645.1; -  
 CC EMBL: X01798; CAA25928.1; -  
 CC EMBL: X01799; CAA25930.1; -  
 CC PIR: A00942; NMSG.  
 CC PDB: 1SGF; 27-MAY-98.  
 CC MEROPS: S01.170; -  
 CC MGD: MGI:97322; Ngf.  
 CC InterPro: IPR001314; Chymotrypsin.  
 CC InterPro: IPR001254; Ser-protease\_Try.  
 CC Pfam: PF00089; trypsin; 1.  
 CC PRINTS: PR00722; CHYMOTRYPSIN.  
 CC SMART: SM00020; Tryp\_Spc; 1.  
 CC PROSITE: PS00240; TRYPsin\_DOM; 1.  
 CC PROSITE: PS00134; TRYPsin\_SER; 1.  
 CC PROSITE: PS00135; TRYPsin\_SER; 1.  
 CC KMW: Hydrolyase; Serine protease; Glycoprotein; Multigene family; zymogen;  
 CC signal; Growth factor; 3D-structure.  
 CC FT SIGNAL 1 18 PROBABLE.  
 CC FT PROPEP 19 24 ACTIVATION PEPTIDE.  
 CC FT CHAIN 25 261 GLANDULAR KALLIKREIN K3.  
 CC FT CHAIN 25 107 NERVE GROWTH FACTOR GAMMA CHAIN 1.  
 CC FT CHAIN 112 261 NERVE GROWTH FACTOR GAMMA CHAIN 2.  
 CC FT ACT\_SITE 65 65 CHARGE RELAY SYSTEM.  
 CC FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM.  
 CC FT ACT\_SITE 213 213 CHARGE RELAY SYSTEM.  
 CC FT DISULFID 31 173  
 CC FT DISULFID 50 66  
 CC FT DISULFID 152 219  
 CC FT DISULFID 184 198  
 CC FT DISULFID 209 234  
 CC FT CARBOHYD 102 102  
 CC FT DOMAIN 25 107 SEGMENT B1.  
 CC FT DOMAIN 112 261 SEGMENT A.  
 CC FT DOMAIN 112 164 SEGMENT C.  
 CC FT DOMAIN 165 261 SEGMENT B2.  
 CC FT CONFLICT 108 111 MISSING (IN REF. 2).  
 CC SQ SEQUENCE 261 AA; 28998 MW; 4870748E174AF7C8 CRC64;

Query Match 47.5%; Score 600.5; DB 1; Length 261;  
 Best Local Similarity 44.7%; Pred. No. 4.7e-48;  
 Matches 106; Conservative 46; Mismatches 78; Indels 7; Gaps 1;

QY 1 RIIFGECKRHSQPMOALFEKTRLLCGATLIPRLMLTAHCLKPRYIVHGOHNLQKE 60  
 DB 24 RVIIGFCEKRNQPMOALFEKTRLLCGATLIPRLMLTAHCLKPRYIVHGOHNLQKE 83  
 QY 61 EGGEOTRTATESPPHGFNNLSL-----PNKDRNDIMLVKASPVSTMAVRPLTSS 113  
 DB 84 EPRNQHVLVSQSPHPHYKPFELMKRNRKRDGDSHNDMLHLISQPDITDGVKVDLPT 143  
 QY 114 RCYTAGTSCILSGWGSSPOLRLPRLTLCANTTIEHOKCENAYPGNTDPMVCAVQOE 173  
 DB 144 EEPKVGSTCLASGWSGTRPKPFTDDICVNLKLNELNECIRAYKEKRYVDIMACAGEMD 203  
 QY 174 GGDSCGDSGGPVLVNCQSLGIIISWGDPACATRRPGVYTKYCKYVDIQTETKNN 230  
 DB 204 GGDTCGDSGGPVLVNCQSLGIIISWGDPACATRRPGVYTKYCKYVDIQTETKNN 260

RESULT 14  
 KLR3\_PRANA

ID	KLRK_PRANA	STANDARD:	PRT:	263 AA.
AC	P32824;			
Dt	01-OCT-1993 (Rel. 27, Created)			
Df	01-OCT-1993 (Rel. 27, Last sequence update)			
Dv	01-NOV-1995 (Rel. 32, Last annotation update)			
DE	Glandular kallikrein, renal precursor (EC 3.4.21.35) (Tissue kallikrein).			
OS	Promys natalensis (African soft-furred rat) (Mastomys natalensis).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mastomys.			
OX	NCBI_TaxID=10112;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-Salivary gland;			
RA	MEDLINE=94226702; PubMed=7909667;			
RT	Fahnestock M.;			
FT	"Characterization of kallikrein cDNAs from the African rodent Mastomys";			
RU	DNA Cell Biol. 13:293-300(1994).			
CC	-I- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS IN KININOGEN TO RELEASE LYS-BRADYKININ.			
CC	-I- CATALYTIC ACTIVITY: Preferential cleavage of Arg-I-Xaa bonds in small molecule substrates. Highly selective action to release kallidin (lysyl-bradykinin) from kininogen involves hydrolysis of Met-I-Xaa or Leu-I-Xaa.			
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/sib.ch).			
CC	--or send an email to license@sib-sib.ch)--			
DR	EMBL: X17352; GAB35232.1; "			
DR	PIR: S15686; S15686.			
DR	HSSP: P00757; ISGF.			
DR	MEROPS: IPR001314; Chymotrypsin.			
DR	InterPro: IPR001254; Ser_protease_Try.			
DR	Pfam: PF00089; trypsin_1			
DR	PRINTS: PR00722; CHYMOTRYPSIN.			
DR	SMART: SM00020; TRYP_Sec: 1.			
DR	PROSITE: PS00240; TRYPSIN_DOM: 1.			
DR	PROSITE: PS00134; TRYPSIN_HIS: 1.			
DR	PROSITE: PS00135; TRYPSIN_SER: 1.			
KW	Hydrolase; Serine protease; Glycoprotein; Multigene family; Zymogen; Signal.			
FT	SIGNAL	1	18	PROBABLE.
FT	PROPEP	19	24	ACTIVATION PEPTIDE (PROBABLE).
FT	CHAIN	25	263	GLANDULAR KALLIKREIN, RENAL.
FT	ACT_SITE	65	65	CHARGE RELAY SYSTEM.
FT	ACT_SITE	121	121	CHARGE RELAY SYSTEM.
FT	ACT_SITE	215	215	CHARGE RELAY SYSTEM.
FT	DISULFD	31	175	BY SIMILARITY.
FT	DISULFD	50	66	BY SIMILARITY.
FT	DISULFD	153	221	BY SIMILARITY.
FT	DISULFD	186	200	BY SIMILARITY.
FT	DISULFD	211	236	BY SIMILARITY.
FT	CARBOND	102	102	N-LINKED (GLCNAC...) (PROBABLE).
SO	SEQUENCE	263 AA;	29130 MW;	AABBQ3BB00337D5 CRC64;
Query Match		47.2%;	Score 596.5;	DB 1; Length 263;
Best Local Similarity		44.8%;	Pred. No. 1.1e-47;	
Matches 107;	Conservative	44;	Mismatches 79;	Indels 9; Gaps 2;
Oy	1 RIIFGECKPHSQMOALFEKTRLGCGATLLAPWLTLAAHCLPKRYIVHLGOHNLOKE 60        ::      ::         :        ::        ::   :			
Dv	24 RIGFGNEKNQSOPAHVAVVRFRAGCGVLDDAWAVFLTAHCYNDKYGMVGKKNNRFED 83        ::      ::         :        ::        ::   :			
Oy	61 EGCEGTATATESPPHPGGNSLPDKH-----RNDIMLVKMASPVSIWAVPPLTS 112			

```
Db      84 EESAOHOLSKLPIPHGFMSLLNDHTHPREDVSNLDMLYRKKRPATITDYVRIDLP 143
Oy      113 SMCVTAGTSLISGWSFS-SPOLRLPHTLRKANITIEHQECENAYPGNITDTMYCASV 171
Db      144 TREPVTSGRCIASGMSGSTPTPEEFEXSHDLCQVYLELLSNEVCANRHTEKVIDFMLCAGE 203
Oy      172 QEGGSDSCGGDSGGFLVNCQSIOGITISMGODPCAITRKGVYTKCYKDWMJOETMKN 230
Db      204 MDGCKDTCVGDSGCLICDGVLGITSMGPFCALPNNGIYTKLIEVRSWIKDVANN 262

RESULT 15
TRY3.CHICK
ID TRY3_CHICK STANDARD; PRY; 248 AA.
AC O90629;
DE 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DN 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin II-P29 precursor (EC 3.4.21.4).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=95251611; PubMed=7733885;
RA Wang K., Gan L., Lee I., Hood L.E.;
RT "Isolation and characterization of the chicken trypsinogen gene family.";
RL Biochem. J. 307:471-479(1995).
CC -1 CATALYTIC ACTIVITY: Preferential cleavage: Arg-I-Xaa, Lys-I-Xaa.
CC -1 SUBCELLULAR LOCATION: Extracellular.
CC -1 TISSUE SPECIFICITY: HIGH LEVELS ARE SEEN IN THE PANCREAS WHILE LOWER LEVELS ARE FOUND IN THE LIVER, SPLEEN AND THYMUS.
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/isb.ch).
or send an email to license@isb-sib.ch).
-----
CC EMBL: U15157; AAA79914.1; -.
DR HSSP: P00763; IDPO.
DR MEROPS: S01.I51; -.
DR InterPro: IPRO001314; Chymotrypsin.
DR InterPro: IPRO01254; Ser_protease_try.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SMO0020; Tryp_SPC; 1.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal; Multipeptide family.
FT SIGNAL 1..16
FT PROPEP 17..25 BY SIMILARITY.
FT CHAIN 26..248 ACTIVATION PEPTIDE (BY SIMILARITY).
FT ACN_SITE 65..65 TRYPSIN II-P29.
FT ACT_SITE 109..109 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 202..202 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 32..162 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 50..66 BY SIMILARITY.
FT DISULFID 134..235 BY SIMILARITY.
FT DISULFID 141..208 BY SIMILARITY.
FT DISULFID 173..187 BY SIMILARITY.
FT DISULFID 198..222 BY SIMILARITY.
FT SITE 196..196 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
```

Seq	SEQUENCE	248 AA	26622 MM	ESL16B07622B588E	CRC64
Query Match		47.2%	Score 596	DB 1	Length 248
Match Local Similarity		49.3%	Pred. No. 1.2e-47		
Matches 112		Conservative 36	Mismatches 73	Indels 6	Gaps 3
QY	1	RIIKGFCEKPHSQPQWALFETKTRILLCGATLIAPFWLITTAACHLPRYIVHLGQHNLOKE 60			
Db	25	KIVGGYTCPEHSPVQYSL-NSGVHFCGSLINSQWVLSAAHCYKSRIQVRLGEVINDVQ 83			
QY	61	EGCEPTKRAATESFPHPGNNLSLPKNDHNDIMLVMAASPVSTITMAVRPLTSSRCVYAGT 120			
Db	84	EDSEYVRSSSVYIRHPKYSITLN---NDMLIKLAAYERSADIDPIALPSSCAKAGT 139			
QY	121	SLGISMGSTSSPOLRLHTLRCAKITITIEHQKCNENAPGNTIDPMWCAVQEGGKDCQ 180			
Db	140	ECLISMGKNTILNSGNYRPELLQCLNAPILLSQEGCEAFPGDITSMIVCGFLFGSKDCQ 199			
QY	181	GDGGPVLVCNOSLAGIISMGDPCAIIRKPGVYTRKCYKRYVDMIOETM 227			
Db	200	GDGGPVLVCNGLQIGIVSMGIG-CALKGYPGVYTRKCYKRYVDMIOETI 245			
RESULT 16					
TRY2_RAT					
ID	TRY2_RAT	STANDARD	PRT	246 AA	
AC	P00763				
DT	21-JUL-1986	(Rel. 01, Created)			
DT	15-JUL-1998	(Rel. 36, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	TRY2				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=85054880; Pubmed=6094547;				
RA	Craik C.S., Choo Q.L., Swift G.H., Quinto C., McDonald R.J.,				
RT	Rutter W.J.;				
RL	"Structure of two related rat pancreatic trypsin genes.";				
RL	J. Biol. Chem. 259:14255-14264(1984).				
RP	[2]				
RP	SEQUENCE OF 9-246 FROM N.A.				
RX	STRAIN=Sprague-Dawley; TISSUE=Pancreas;				
RC	MEDLINE=82265624; Pubmed=6896710;				
RA	McDonald R.J., Stary S.J., Swift G.H.;				
RT	"Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide				
RL	sequences of the cloned cDNAs.";				
RL	J. Biol. Chem. 257:9724-9732(1982).				
RN	[3]				
RP	X-RAY CRYSTALLOGRAPHY (1.59 ANGSTROMS).				
RX	MEDLINE=91351998; Pubmed=1881877;				
RA	Earnest T., Fauman E., Craik C.S., Stroud R.;				
RT	"1.59-A structure of trypsin at 120 K: comparison of low temperature				
RL	and room temperature structures.";				
RL	Proteins 10:171-187(1991).				
RN	[4]				
RP	X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).				
RX	MEDLINE=96214506; Pubmed=6634241;				
RA	Brihen L.S., Willett W.S., Craik C.S., Fletcher R.J.;				
RT	"X-ray structures of a designed binding site in trypsin show metal-				
RL	dependent geometry.";				
CC	biochemistry 35:5999-6009(1996).				
CC	-1- CATALYTIC ACTIVITY: Preferential cleavage: Arg- -Xaa, Lys- -Xaa.				
CC	-1- SUBCELLULAR LOCATION: Extracellular.				
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				

```

CC modified and this statement is not removed. Usage by end for commercial/
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL; V01274; CAA24581.1; -.
DR EMBL; L00131; AAA98517.1; -.
DR EMBL; L00130; AAA98517.1; JOINED.
DR PIR; A22657; TRRT2.
DR PDB; 1ANB; 01-APR-97.
DR PDB; 1ANC; 01-APR-97.
DR PDB; 1ANE; 01-APR-97.
DR PDB; 1AME; 24-DEC-97.
DR PDB; 1SIU; 11-JUL-96.
DR PDB; 1SIW; 11-JUL-96.
DR PDB; 1SLX; 11-JUL-96.
DR PDB; 1DPO; 07-JUL-97.
DR PDB; 3TGI; 23-DEC-98.
DR PDB; 3TGI; 23-DEC-98.
DR MEROPS; S01.258; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR HydroLase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Multigene family; 3d-structure.
FT SIGNAL 1 15
FT PROPEP 16 23 ACTIVATION PEPTIDE.
FT CHAIN 24 246 TRYPSIN II, ANIONIC.
FT ACT_SITE 63 63 CHANGE RELAY SYSTEM.
FT ACT_SITE 107 107 CHANGE RELAY SYSTEM.
FT ACT_SITE 200 200 CHANGE RELAY SYSTEM.
FT DISULFID 30 160
FT DISULFID 48 64
FT DISULFID 132 233
FT DISULFID 139 206
FT DISULFID 171 185
FT DISULFID 196 220
FT SITE 194 194
FT CONFLICT 84 84 REQUIRED FOR SPECIFICITY. (BY SIMILARITY).
FT CONFLICT 88 88 N -> D (IN REF. 1).
FT CONFLICT 88 88 V -> I (IN REF. 1).
SQ SEQUENCE 246 AA; 26228 MW; A8D3630809AE8606 CRC64;

Query Match 47.08; Score 593; DB 1; Length 246;
Best Local Similarity 49.18; Pred. No. 2.2e-47;
Matches 113; Conservative 33; Mismatches 78; Indels 6; Gaps 3;

1 RIKKEFEKPSQHPQALAEKFEKTRLCGATLIAPMLTAAMCLKPRIVHIGQHNLOKE 60
23 KIVGVTQIGNSVPRQVSL-NSGTHFCGSLINDQWVYSAHCHYRSQVRLGEININVL 81
61 EGCEQRTATSPFPHGFGNNSLPKNDHRNDIMLVKMAVPSVITMAVRPLTSSRCVTAGT 120
82 EGNEFVMAAKTIKHPNDKRTLN---NDIMLILSSPVKLNARVATVLPSSCAPAGT 137
121 SCLISGMSSTSPQRLRHLHTRCANITITIEHQKCNENAPGNTDPMNCASVQEGKDSQ 180
138 QCLISGMSNTLSSGVNEFDLIQCLDAPLIPQADCCASTPGKLTDMVAVGVLEGKDSQ 197
181 GDSGPIVACNOSLOGIISWQDPCATITRKPGVYTVKVCYVVMIOETMRNN 230
198 GDSGPIVACNOSLOGIISWQDPCATITRKPGVYTVKVCYVVMIOETMRNN 246

Search completed: December 30, 2002, 14:50:22
Job time : 13 secs

```

---





**THIS PAGE BLANK (USPTO)**